

# Cointegrated Continuous-time Linear State Space and MCARMA Models

Vicky Fasen-Hartmann <sup>†‡</sup>

Markus Scholz <sup>†</sup>

In this paper we define and characterize cointegrated continuous-time linear state-space models. A main result is that a cointegrated continuous-time linear state-space model can be represented as a sum of a Lévy process and a stationary linear state-space model. Moreover, we prove that the class of cointegrated multivariate Lévy-driven autoregressive moving-average (MCARMA) processes, the continuous-time analogues of the classical vector ARMA processes, is equivalent to the class of cointegrated continuous-time linear state space models. Necessary and sufficient conditions for MCARMA processes to be cointegrated are given as well extending the results of Comte [7] for MCAR processes. The conditions depend on the autoregressive polynomial. Finally, we investigate cointegrated continuous-time linear state-space models observed on a discrete time-grid and derive an error correction form for this model. The error correction form is based on an infinite linear filter in contrast to the finite linear filter for VAR models.

*JEL Classification:* C10, C32, C51

*Keywords:* Canonical form, cointegration, error correction form, Granger representation theorem, Kalman filter, MCARMA process, state-space model.

## 1 Introduction

Many time series do not behave in a stationary way, e.g. financial time series. Hence, non-stationary models are of particular interest in order to model such behavior. One class of non-stationary processes are cointegrated processes. A  $d$ -dimensional stochastic process  $(Y(t))_{t \geq 0}$  without deterministic component is called *integrated* if  $(Y(t))_{t \geq 0}$  is non-stationary but has stationary increments. If additionally there exists a vector  $\beta \neq 0$ ,  $\beta \in \mathbb{R}^d$ , such that  $(\beta^\top Y(t))_{t \geq 0}$  is stationary then  $(Y(t))_{t \geq 0}$  is called *cointegrated* with cointegration vector  $\beta$ . The number of linear independent cointegration relations is called *cointegration rank* and the space spanned by all linear independent cointegration vectors is called *cointegrating space*. Examples of cointegrated time series are, e.g., exchange rates, foreign currency spot and futures/forwards rates, stock prices within an industry and interest rates in different countries (cf. Brenner and Kroner [4] and references therein). It was Clive Granger who showed that statistical inference of such non-stationary time series with the classical stationary methodology can lead to inadequate results. The seminal works by Granger [14] in 1981 and Engle and Granger [9] in 1987

<sup>†</sup>Institute of Stochastics, Englerstraße 2, D-76131 Karlsruhe, Germany.

<sup>‡</sup>Financial support by the Deutsche Forschungsgemeinschaft through the research grant FA 809/2-2 is gratefully acknowledged.

lay the foundation for the field of cointegration analysis in discrete time. One of the most important results is the Granger representation theorem which connects the moving average, autoregressive and error correction representations of cointegrated time series. Johansen [16, 17] characterizes cointegration for VAR (vector autoregressive) processes precisely by making assumptions on the autoregressive polynomial in the Johansen-Granger representation theorem.

In this paper we study cointegrated continuous-time linear stochastic state-space models and cointegrated multivariate continuous-time ARMA( $p, q$ ) (MCARMA( $p, q$ )) processes, respectively which are the natural extension of VARMA (vector autoregressive moving average) processes from discrete to continuous time. It is well known that the class of stationary linear state-space models and the class of stationary MCARMA processes are equivalent; see [27, Corollary 3.4]. Both model classes are driven by a Lévy process. An  $\mathbb{R}^m$ -valued stochastic process  $L = (L(t))_{t \geq 0}$  is a Lévy process if  $L(0) = 0_m$   $\mathbb{P}$ -a.s., it has stationary and independent increments and càdlàg sample paths; see the excellent monograph of Sato [26] for details on Lévy processes. Throughout this paper we will assume

**Assumption A.** *The Lévy process  $L$  satisfies  $\mathbb{E}L(1) = 0_m$  and  $\mathbb{E}\|L(1)\|^2 < \infty$  and the covariance matrix  $\Sigma_L = \mathbb{E}[L(1)L(1)^T]$  is non-singular.*

Then an  $\mathbb{R}^d$ -valued *continuous-time linear state-space model*  $(A, B, C, L)$  of dimension  $N \in \mathbb{N}$  is characterized by an  $\mathbb{R}^m$ -valued Lévy process, a transition matrix  $A \in \mathbb{R}^{N \times N}$ , an input matrix  $B \in \mathbb{R}^{N \times m}$  and an observation matrix  $C \in \mathbb{R}^{d \times N}$ . It consists of the state equation

$$dX(t) = AX(t)dt + BdL(t)$$

and the observation equation

$$Y(t) = CX(t) \quad \text{for } t \geq t_0 \geq 0. \quad (1.1)$$

The state vector process  $(X(t))_{t \geq t_0}$  is an  $\mathbb{R}^N$ -valued process and the output process  $(Y(t))_{t \geq t_0}$  is  $\mathbb{R}^d$ -valued. Every solution of (1.1) has the representation

$$Y(t) = C \exp(A(t - t_0))X(t_0) + C \int_{t_0}^t \exp(A(t - u))B dL(u).$$

A solution  $Y$  is called *causal*, if for all  $t \geq t_0$ ,  $Y(t)$  is independent of the  $\sigma$ -algebra generated by  $\{L(s) : s > t\}$ .

On the other hand, the idea behind the definition of an  $\mathbb{R}^d$ -valued MCARMA( $p, q$ ) process ( $p > q$  positive integers) is that it is the solution to the differential equation

$$P(D)Y(t) = Q(D)DL(t) \quad \text{for } t \geq t_0, \quad (1.2)$$

where  $D$  is the differential operator with respect to  $t$ ,

$$P(z) := I_{d \times d}z^p + P_1z^{p-1} + \dots + P_{p-1}z + P_p \quad (1.3)$$

is the autoregressive polynomial with  $P_1, \dots, P_p \in \mathbb{R}^{d \times d}$  and

$$Q(z) := Q_0z^q + Q_1z^{q-1} + \dots + Q_{q-1}z + Q_q \quad (1.4)$$

is the moving average polynomial with  $Q_0, \dots, Q_q \in \mathbb{R}^{d \times m}$ . However, since a Lévy process is not differentiable the formal definition of an MCARMA process is given later but it is a special linear

state space model.

The aim of this paper is to characterize continuous-time cointegrated linear state-space and MCARMA models and to relate both model classes. Cointegration in continuous time started being of interest in the early 1990s with Phillips [23]. In this work, Phillips investigated stochastic differential equations driven by a differentiable stationary process. The connection between cointegrated discrete-time models and continuous-time models were analyzed by Chambers [5]. The literature on Gaussian MCAR(1,0) processes is rich, e.g., [19, 20, 31]. One of the first going away from the Gaussian assumption and the order (1,0) was Comte [7]; he derived a characterization of continuous-time integrated and cointegrated processes, and in particular, he presented an error correction form and a characterization of cointegration for MCARMA( $p,0$ ) processes. The processes considered in Fasen [10, 11] are special cases of cointegrated MCARMA processes.

The paper is structured on the following way. In Section 2 we introduce cointegrated linear state-space models in continuous-time and show that the definition is well-defined. An important conclusion is that a cointegrated linear state-space process has a representation as a sum of a non-stationary and a stationary process; the non-stationary process is a Lévy process and the stationary process is a stationary linear state-space model. This characterization can be used as definition of a cointegrated state-space model as well. Moreover, we investigate cointegrated MCARMA processes. An important result is that the class of cointegrated MCARMA models and the class of cointegrated linear state-space models are equivalent completing [27, Corollary 3.4] to the non-stationary case.

Probabilistic properties of cointegrated linear state-space models are content of Section 3. First, we derive an alternative characterization of cointegration for MCARMA processes extending the results of Comte [7] for MCAR processes. The property of cointegration of an MCARMA process is given in the matrices  $P_p$  and  $P_{p-1}$  of the autoregressive polynomial (1.3). Furthermore, we investigate cointegrated linear state-space models sampled at a discrete-time grid because they are of particular interest in high-frequency data.

In the last section, Section 4, of this paper we derive an error correction form for a cointegrated linear state-space model sampled at a discrete time-grid. The error correction form enables us to analyze the short-run as well as the long-run behaviour; it is the first step to develop a parameter estimation method for a cointegrated model (cf. Fasen and Scholz [12]). However, in order to obtain an error correction form for the sampled process we need to calculate the so-called linear innovations. This is done with the Kalman filter whose name dates back to Rudolf E. Kalman [18]. The derived error correction form is similar to the original error correction form presented by Engle and Granger [9] for VAR models. The main difference is that we have an infinite order linear filter instead of a finite linear filter. We show that the cointegration information is contained in parts of the filter and is thus not lost by sampling and filtering. Furthermore, the error correction form is used to derive probabilistic properties of the linear innovations.

## Notation

We use as norms the Euclidean norm  $\|\cdot\|$  in  $\mathbb{R}^d$  and the Frobenius norm  $\|\cdot\|$  for matrices, which is submultiplicative.  $\Re(z)$  denotes the real part of a complex number  $z \in \mathbb{C}$ . The matrix  $0_{d \times s}$  is the zero matrix in  $\mathbb{R}^{d \times s}$  and  $I_d$  is the identity matrix in  $\mathbb{R}^{d \times d}$ . For a matrix  $A \in \mathbb{R}^{d \times d}$  we denote by  $A^\top$  its transpose, by  $\det(A)$  its determinant, by  $\text{rank} A$  its rank and by  $\lambda_{\max}(A)$  its largest eigenvalue. For a matrix  $A \in \mathbb{R}^{d \times s}$  with  $\text{rank} A = s$ ,  $A^\perp$  is an  $d \times (d-s)$  matrix with  $\text{rank}(d-s)$  satisfying  $A^\top A^\perp = 0_{s \times (d-s)}$  and  $A^\perp A = 0_{(d-s) \times s}$ . For two matrices  $A, B$  we write  $\text{diag}(A, B)$  for a block diagonal matrix whose first block is the matrix  $A$  and the second block is the matrix  $B$ . The space of all  $m \times n$  real-valued matrices is denoted with  $M_{m,n}(\mathbb{R})$ , the set of  $m$ -dimensional symmetric positive-definite matrices is denoted by  $\mathbb{S}_m^{++}$  and  $GL_N(\mathbb{R}) := \{A \in \mathbb{R}^{N \times N} : \det A \neq 0\}$  for some  $N \in \mathbb{N}$ .

## 2 Cointegrated state-space models and cointegrated MCARMA models

### 2.1 Cointegrated state-space models

First, we present definitions related to linear state-space models which we need subsequently. These definitions enable us to imply restrictions on the state-space model in order to achieve uniqueness in the output process and to define a cointegrated model.

**Definition 2.1.**

(a) The continuous-time linear state-space model is **observable** if the observability matrix

$$\mathcal{O}_{CA} := \begin{pmatrix} C^\top & (CA)^\top & \dots & (CA^{N-1})^\top \end{pmatrix}^\top \in M_{dN,N}(\mathbb{R})$$

has full rank, i.e. if  $\text{rank}(\mathcal{O}_{CA}) = N$ .

(b) The continuous-time linear state-space model is **controllable** if the controllability matrix

$$\mathcal{C}_{AB} := \begin{pmatrix} B & AB & \dots & A^{N-1}B \end{pmatrix} \in M_{N,mN}(\mathbb{R})$$

has full rank, i.e. if  $\text{rank}(\mathcal{C}_{AB}) = N$ .

Another desired property of linear state-space models is a minimal dimension because otherwise there exists state-space models of smaller dimension which generate the same output process.

**Definition 2.2.** Let  $A \in M_N(\mathbb{R})$ ,  $B \in M_{N,m}(\mathbb{R})$  and  $C \in M_{d,N}(\mathbb{R})$ . The matrix triple  $(A, B, C)$  is called an **algebraic realization** of a rational matrix function  $k \in M_{d,m}(\mathbb{R}\{z\})$  of dimension  $N$  if  $k(z) = C(zI_N - A)^{-1}B$ . The function  $k : z \mapsto C(zI_N - A)^{-1}B$  is called **transfer function** of the state-space model  $(A, B, C)$ . The triplet  $(A, B, C)$  is called **minimal** if there exists no other algebraic realization with dimension smaller than  $N$ . The dimension of a minimal realization of  $k$  is called **McMillan degree** of  $k$ .

Thus, non-minimality is a source of non-uniqueness of the state-space model. Minimality guarantees that we consider only components of the state vector which are relevant for the output process. If we have a non-minimal system there might be non-stationary components having no effect on the output process. Therefore, this property implies a one-to-one correspondence of the non-stationarity of the state process and the output process. A state-space model is minimal if and only if its controllable and observable (cf. Hannan and Deistler [15, Theorem 2.3.3]). Last but not least, we give the formal definition of observational equivalence of state-space models.

**Definition 2.3.** A minimal linear state-space model  $(A, B, C)$  is called **observationally equivalent** to the minimal model  $(\tilde{A}, \tilde{B}, \tilde{C})$  if they give rise to the same transfer function.

Hence,  $(A, B, C)$  and  $(\tilde{A}, \tilde{B}, \tilde{C})$  are observationally equivalent if and only if there exists a nonsingular transformation matrix  $T \in GL_N(\mathbb{R})$  such that  $A = T\tilde{A}T^{-1}$ ,  $B = T\tilde{B}$  and  $C = \tilde{C}T^{-1}$ . Such a transformation leads to a corresponding basis change of the state vector to  $\tilde{X}(t) = TX(t)$ .

The aim is to define a cointegrated linear state-space model. For this purpose we introduce a convenient canonical form for a state-space model. A canonical form is a unique representation of the class of observationally equivalent linear state-space models. An example for a canonical form

is the echelon canonical form (see e.g. Schlemm and Stelzer [28, Section 4.1]). Our canonical form for the cointegrated continuous-time linear state-space model has a similar structure as the canonical form for the discrete-time linear state-space model presented in Bauer and Wagner [1, Theorem 2 and Theorem 3]. The advantage of this canonical form is that the non-stationary and stationary part are decoupled and can be transformed separately. Moreover, this form enables us to use existing results for stationary state-space models and Lévy processes in the following. Before we can state the result we repeat the definition of a positive lower triangular matrix: a matrix  $M \in M_{d,c}(\mathbb{R})$  is positive lower triangular if  $M$  is lower triangular and the first non-zero entry in each row is positive.

**Theorem 2.4.** *Let  $(A, B, C, L)$  be a  $d$ -dimensional minimal state-space model which satisfies  $\sigma(A) \subset \{(-\infty, 0) + i\mathbb{R}\} \cup \{0\}$  and the algebraic and the geometric multiplicity of the eigenvalue zero is equal to  $c$ ,  $0 \leq c \leq d$ . Then there exists a unique observationally equivalent minimal state-space representation given by*

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \end{pmatrix} = \begin{pmatrix} 0_{c \times c} & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & A_2 \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} + \begin{pmatrix} B_1 dL(t) \\ B_2 dL(t) \end{pmatrix}, \quad (2.1)$$

$$Y(t) = C_1 X_1(t) + C_2 X_2(t), \quad t \geq t_0,$$

where

- (i)  $|\lambda_{\max}(\exp(A_2))| < 1$ ,
- (ii) the matrix  $C_1 \in M_{d,c}(\mathbb{R})$  satisfies  $C_1^\top C_1 = I_c$  and  $C_1$  is a positive lower triangular matrix,
- (iii) the matrix  $B_1 \in M_{c,d}(\mathbb{R})$  (is not restricted),
- (iv) the stationary state-space model  $(A_2, B_2, C_2)$  is real-valued and given in canonical form.

Moreover,  $Y$  is a sum of an initial value, a Lévy process and a stationary state-space model

$$Y(t) = C_1 X_1(t_0) + C_1 B_1 L(t) + C_2 \int_{-\infty}^t \exp(A_2(t-u)) B_2 dL(u), \quad t \geq t_0 \quad (2.2)$$

if we choose  $X_2(t_0)$  appropriately.

*Proof.* Solving (2.1) leads directly to

$$Y(t) = C_1 X_1(t_0) + C_2 \exp(A_2(t-t_0)) X_2(t_0) + C_1 B_1 L(t) + C_2 \int_{t_0}^t \exp(A_2(t-u)) B_2 dL(u),$$

which in the end gives (2.2).

We define the following matrices

$$A^* := \begin{pmatrix} 0_{c \times c} & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & A_2 \end{pmatrix}, \quad B^* := \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \quad \text{and} \quad C^* := (C_1 \quad C_2). \quad (2.3)$$

For the existence of the representation (2.1) we need to show that there exists a  $T \in GL_N(\mathbb{R})$  which transforms the state-space model  $(A, B, C)$  to the desirable form  $(A^*, B^*, C^*)$  and satisfies all restrictions (i)-(iv). Afterwards, we have to show that this transformation matrix is unique which results in the uniqueness of this representation.

**Existence:** Due to the eigenvalue assumption on the matrix  $A$  the upper part of  $A^*$  is just the Jordan normal form corresponding to the eigenvalue zero. Thus there exists a transformation matrix  $T'$  such that

$$T'AT'^{-1} = \begin{pmatrix} 0_{c \times c} & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & A'_2 \end{pmatrix} =: A',$$

where the eigenvalues of  $A'_2$  coincide with the non-zero eigenvalues of  $A$  which have by assumption strictly negative real parts. Otherwise  $A'_2$  is not specified yet. Further,

$$B' := T'B = \begin{pmatrix} B_1^\top & B_2^\top \end{pmatrix}^\top \quad \text{and} \quad C' := CT'^{-1} = \begin{pmatrix} C_1 & C_2 \end{pmatrix}.$$

Since the block-diagonal structure of  $A'$  is preserved by block transformations, we consider in the following only block-diagonal transformation matrices  $T'' = \text{diag}(T_1'', T_2'')$  (see Gantmacher [13], p.231) resulting in

$$\begin{aligned} A'' &:= T''A'T''^{-1} = \begin{pmatrix} 0_{c \times c} & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & T_2''A'_2T_2''^{-1} \end{pmatrix}, \quad B'' := T''B' = \begin{pmatrix} T_1''B_1' \\ T_2''B_2' \end{pmatrix} \quad \text{and} \\ C'' &:= C'T''^{-1} = \begin{pmatrix} C_1'T_1''^{-1} & C_2'T_1''^{-1} \end{pmatrix}. \end{aligned}$$

Obviously, there exists a transformation matrix  $T_1''$  such that  $C_1 := C_1'T_1''^{-1}$  satisfies (ii) and  $B_1 := T_1''B_1'$ . Since the eigenvalues of  $A'_2$  have strictly negative real parts,  $(A'_2, B'_2, C'_2)$  forms a stationary linear state-space model. Hence, there exists a transformation matrix  $T_2''$  such that  $A_2 := T_2''A'_2T_2''^{-1}$ ,  $B_2 := T_2''B'_2$  and  $C_2 := C_2'T_2''^{-1}$  satisfies (iv). Moreover, the eigenvalues of  $A'_2$  and hence,  $A_2$  have strictly negative real parts so that (i) is satisfied as well. Finally, we set  $T = T''T'$  and  $(A^*, B^*, C^*) = (A'', B'', C'')$ .

**Uniqueness:** Assume there exists matrices

$$\tilde{A} := \begin{pmatrix} 0_{c \times c} & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & \tilde{A}_2 \end{pmatrix}, \quad \tilde{B} := \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{pmatrix} \quad \text{and} \quad \tilde{C} := \begin{pmatrix} \tilde{C}_1 & \tilde{C}_2 \end{pmatrix},$$

so that the state-space model  $(\tilde{A}, \tilde{B}, \tilde{C})$  satisfies the assumptions of this theorem as well. But then there exists a block diagonal transformation  $T = \text{diag}(T_1, T_2)$  with  $(\tilde{A}, \tilde{B}, \tilde{C}) = (TA^*T^{-1}, TB^*, C^*T^{-1})$ . To be more precise  $\tilde{A}_2 = T_2A_2T_2^{-1}$ ,  $\tilde{B}_1 = T_1B_1$ ,  $\tilde{B}_2 = T_2B_2$ ,  $\tilde{C}_1 = C_1T_1^{-1}$  and  $\tilde{C}_2 = C_2T_2^{-1}$ . Since  $(A_2, B_2, C_2)$  and  $(\tilde{A}_2, \tilde{B}_2, \tilde{C}_2)$ , respectively are given in canonical form (cf. restriction (iv)),  $T_2$  has to be the identity matrix. In order to prove uniqueness it remains to show that  $T_1$  is the identity matrix as well. Due to  $\tilde{C}_1^\top \tilde{C}_1 = I_c$  by (ii), we obtain

$$I_c = \tilde{C}_1^\top \tilde{C}_1 = (C_1T_1)^\top C_1T_1 = T_1^\top C_1^\top C_1T_1 = T_1^\top T_1$$

and thus,  $T_1$  has to be orthogonal. If we now exploit the fact that  $C_1$  and  $\tilde{C}_1$  are both lower triangular matrices, we further get that  $T_1$  is a lower triangular matrix itself and therefore the entries  $t_{ij}$ , for  $i < j$ , must be zero or otherwise the lower triangular structure would not be preserved. Real-valued orthogonal matrices have only eigenvalues 1 or  $-1$  and thus, the diagonal form of the transformation matrix  $T_1$  consists only of the entries 1 or  $-1$ . Next, we utilize the restriction that the first non-zero element  $c_{jji}$  in each column of  $C_1$  is positive. We start with the last column where the first positive entry  $c_{jcc}$  is multiplied by  $t_{cc}$ . This product must be positive and hence,  $t_{cc}$  must be positive. Since



the diagonal entries of a triangular quadratic matrix are the eigenvalues itself this implies that  $t_{cc} = 1$ . Thus, all columns of  $T_1$  are orthonormal. The  $(c-1)$ -th column has only two entries which are non-zero and in order to be orthonormal to the  $c$ -th unit vector it must be a unit vector itself, i.e. it must be  $e_{c-1}$ . Iterating this procedure leads to  $T_1 = I_c$  and consequently we have a unique form for the state-space model under the stated restrictions.  $\square$

For the rest of the paper, we abbreviate the stationary part with

$$Y_2(t) := C_2 \int_{-\infty}^t \exp(A_2(t-u)) B_2 dL(u), \quad t \geq t_0. \quad (2.4)$$

We show in the next lemma that the assumption of minimality of the state-space model  $(A, B, C, L)$  is equivalent to the assumption of minimality of the stationary state-space model  $(A_2, B_2, C_2, L)$  combined with assumptions on the matrices  $B_1$  and  $C_1$ . These assumptions on the decoupled model are often easier to verify than the assumption of the minimality of the model  $(A, B, C, L)$  itself.

**Lemma 2.5.** *Let  $(A, B, C, L)$  be a  $d$ -dimensional linear state-space model with block representation of  $A, B, C$  as given in (2.3),  $\sigma(A_2) \subset \{(-\infty, 0) + i\mathbb{R}\}$  and  $0 \leq c \leq d$ . Then  $(A, B, C)$  is minimal if and only if  $B_1$  has full row rank and  $C_1$  has full column rank (i.e.  $\text{rank } B_1 = \text{rank } C_1 = c$ ), and the representation  $(A_2, B_2, C_2)$  of the stationary subsystem is minimal.*

*Proof.* Minimality is equivalent to the condition that the controllability and observability matrices  $\mathcal{C}_{AB}$  and  $\mathcal{O}_{AC}$ , respectively have full rank. We prove an alternative criterion for observability (cf. Bernstein [2, Proposition 12.3.13]). Therefore, for all eigenvalues  $\lambda$  of  $A$  we have to determine the rank of

$$\begin{pmatrix} \lambda I_c & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & \lambda I_{N-c} - A_2 \\ C_1 & C_2 \end{pmatrix}.$$

We consider two cases, beginning with the eigenvalue  $\lambda = 0$  which simplifies the matrix to

$$\begin{aligned} \text{rank} \begin{pmatrix} 0_c & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & -A_2 \\ C_1 & C_2 \end{pmatrix} &= \text{rank} \begin{pmatrix} -A_2 & 0_{(N-c) \times c} \\ C_2 & C_1 \end{pmatrix} = (N-c) + \text{rank}(C_1 + C_2 A_2^{-1} 0_{(N-c) \times c}) \\ &= (N-c) + c = N. \end{aligned}$$

The last equations follow by Bernstein [2, Proposition 2.8.3], the fact that  $C_1$  has full rank  $c$  and the invertibility of  $A_2$ . In the case  $\lambda \neq 0$  we get

$$\begin{aligned} \text{rank} \begin{pmatrix} \lambda I_c & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & \lambda I_{N-c} - A_2 \\ C_1 & C_2 \end{pmatrix} &= c + \text{rank} \left( \begin{pmatrix} \lambda I_{N-c} - A_2 \\ C_2 \end{pmatrix} - \begin{pmatrix} 0_{(N-c) \times c} \\ C_1 \end{pmatrix} \cdot \lambda^{-1} I_c \cdot 0_{c \times (N-c)} \right) \\ &= c + \text{rank} \begin{pmatrix} \lambda I_c - A_2 \\ C_2 \end{pmatrix} = N. \end{aligned}$$

Again, we used Bernstein [2, Proposition 2.8.3] in combination with the fact that the stationary part is minimal by assumption and hence, observable. As a consequence, we have that  $\text{rank}(\mathcal{O}_{CA}) = N$  and the model is observable. Analogously, we obtain  $\text{rank}(\mathcal{C}_{AB}) = N$ . Together with the observability this gives the minimality of the model.  $\square$

Although in Theorem 2.4 (iii) we do not assume any specific restriction on  $B_1$ , the assumption on the minimality of the state-space model implies that  $B_1$  has full rank  $c$ . Eventually, we have to check whether the model is indeed cointegrated.

**Lemma 2.6.** *Let  $Y$  be given as in Theorem 2.4 with  $0 < c < d$ . Then  $Y$  is cointegrated with cointegration space spanned by  $C_1^\perp$  and cointegration rank  $d - c = \text{rank } C_1^\perp$ . In particular,  $C_1^\perp Y$  is stationary.*

*Proof.* A conclusion of Lemma 2.5 is that  $C_1 B_1 \neq 0_{d \times d}$  so that  $Y$  as given in (2.2) is indeed an integrated process since the Lévy process  $(C_1 B_1 L(t))_{t \geq 0}$  is a non-stationary process with strictly stationary increments. Moreover,  $C_1^\perp Y = C_1^\perp Y_2$  is a stationary process. Since  $C_1^\top C_1 = I_c$  due to Theorem 2.4 (ii),  $B_1$  has full rank  $c$  due to Lemma 2.5 and  $\Sigma_L$  is non-singular due to Assumption A, every component of  $C_1^\top Y$  is non-stationary. Hence,  $C_1^\perp$  spans the cointegration space with  $\text{rank } C_1^\perp = d - c$ .  $\square$

From these considerations the following definition is well-defined.

**Definition 2.7.** *Let  $(A, B, C, L)$  be a  $d$ -dimensional minimal state-space model of dimension  $N$  which satisfies  $\sigma(A) \subset \{(-\infty, 0) + i\mathbb{R}\} \cup \{0\}$  and the algebraic and the geometric multiplicity of the eigenvalue zero is equal to  $c$  with  $0 < c < d$ . Then the output process  $Y$  is called **cointegrated continuous-time linear state-space model**.*

As we already mentioned, we can alternatively define the cointegrated linear state-space model as a sum of a Lévy process and a stationary linear state-space model which is verified in the next corollary.

**Corollary 2.8.** *The following equivalences hold:*

- (i)  $Y$  is a cointegrated linear state-space model.
- (ii)  $Y$  has the representation

$$Y(t) = C_1 X_1(t_0) + C_1 B_1 L(t) + C_2 \int_{-\infty}^t \exp(A_2(t-u)) B_2 dL(u), \quad t \geq t_0,$$

where  $C_1 \in M_{d,c}(\mathbb{R})$ ,  $B_1 \in M_{c,m}(\mathbb{R})$  with  $\text{rank } C_1 = \text{rank } B_1 = c$  and  $\sigma(A_2) \subset (-\infty, 0) + i\mathbb{R}$ .

*Proof.* The result is an immediate consequence of Theorem 2.4 and Lemma 2.6. Note that the model  $(A_2, B_2, C_2)$  needs not to be minimal. In this case, we find a observationally equivalent minimal model and thus, obtain a cointegrated linear state-space model.  $\square$

For this reason we also write  $(A_2, B_1, B_2, C_1, C_2, L)$  for a cointegrated linear state-space model.

**Corollary 2.9.** *The cointegrated state-space model  $Y$  is causal.*

*Proof.* This is obvious due to representation (2.2).  $\square$

## 2.2 Cointegrated MCARMA process

First, we present the formal definition of an MCARMA process.



**Definition 2.10.** Define

$$A = \begin{pmatrix} 0_{d \times d} & I_{d \times d} & 0_{d \times d} & \cdots & 0_{d \times d} \\ 0_{d \times d} & 0_{d \times d} & I_{d \times d} & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0_{d \times d} \\ 0_{d \times d} & \cdots & \cdots & 0_{d \times d} & I_{d \times d} \\ -P_p & -P_{p-1} & \cdots & \cdots & -P_1 \end{pmatrix} \in \mathbb{R}^{pd \times pd},$$

$C = (I_{d \times d}, 0_{d \times d}, \dots, 0_{d \times d}) \in \mathbb{R}^{d \times pd}$  and  $B = (\beta_1^\top \cdots \beta_p^\top)^\top \in \mathbb{R}^{pd \times m}$  with

$$\beta_1 := \dots := \beta_{p-q-1} := 0_{d \times m} \quad \text{and} \quad \beta_{p-j} := - \sum_{i=1}^{p-j-1} P_i \beta_{p-j-i} + Q_{q-j}, \quad j = 0, \dots, q.$$

Then the  $\mathbb{R}^d$ -valued **MCARMA**( $p, q$ ) **process**  $Y = (Y(t))_{t \geq t_0}$  is defined by the state-space equation

$$Y(t) = CX(t) \quad \text{for } t \geq t_0, \quad (2.5)$$

where  $X$  is the solution to the  $pd$ -dimensional stochastic differential equation

$$dX(t) = AX(t) dt + B dL(t). \quad (2.6)$$

If  $\sigma(A) \subseteq (-\infty, 0) + i\mathbb{R}$  then  $Y$  is called **stationary MCARMA process**, and if  $\sigma(A) \subset \{(-\infty, 0) + i\mathbb{R}\} \cup \{0\}$  where the algebraic and the geometric multiplicity of the eigenvalue zero is equal to  $c$  with  $0 < c < d$  and the state-space model  $(A, B, C)$  is minimal then  $Y$  is called **cointegrated MCARMA process**.

**Remark 2.11.**

- (i) In particular, MCARMA(1,0) processes and the process  $X$  in (2.6) are multivariate Ornstein-Uhlenbeck processes.
- (ii) It is obvious that an MCARMA process is a special state-space model and by Definition 2.7 that a cointegrated MCARMA process is a cointegrated state-space model. Due to Marquardt and Stelzer [22, Lemma 3.8] we already know that  $0 \in \sigma(A)$  iff  $0 \in \sigma(P_p)$ . Thus the matrix  $P_p$  plays a crucial role in the characterization of cointegration.
- (iii) In the case of stationary MCARMA processes Marquardt and Stelzer [22] give details on the well-definedness of this model and that the MCARMA process can be interpreted as the solution of (1.2). Moreover, Schlemm and Stelzer [27, Corollary 3.4] prove that the class of stationary MCARMA processes and the class of stationary state-space models are equivalent.
- (iv) By rewriting (2.5) and (2.6) line by line it is possible to see that an MCARMA process is the solution to (1.2).

The question which arises now is what kind of relation is between the class of cointegrated state space and cointegrated MCARMA models. The following lemma will answer this.

**Lemma 2.12.** The class of cointegrated MCARMA( $p, q$ ) processes and cointegrated state-space models with  $N = pd$  are equivalent.

*Proof.* By definition, every cointegrated MCARMA process is a cointegrated state-space model. Conversely, given a cointegrated state-space model  $Y$ , it can be transformed into the form of Definition 2.10 using as transformation matrix the observability matrix  $T = \mathcal{O}_{CA}$ , see Shieh et al. [30, Theorem 3.3].  $\square$

### 3 Properties of cointegrated MCARMA processes

#### 3.1 Characterization of cointegrated MCARMA processes

An important representation of cointegrated processes in discrete time is the so-called error correction form which was first introduced by Sargan [25]. We can define an analog version for the continuous-time case.

**Definition 3.1.** *Let  $Y$  be a cointegrated MCARMA process as solution of (1.2). The **error correction form** in continuous time is given by*

$$P^*(D)DY(t) = -P_p Y(t) + Q(D)DL(t), \quad (3.1)$$

where the polynomial  $P^*$  has the representation  $P^*(z) := \frac{P(z) - P_p}{z}$  and  $DY$  is the differential of  $Y$  in the  $L^2$  sense.

If the process has a first-order mean square derivative  $DY$ , the definition of cointegration is equivalent to  $DY$  to be stationary due to [7, Proposition 1].

The following result characterizes cointegration with respect to the matrices  $P_p$  and  $P_{p-1}$  in the autoregressive polynomial; it is an extension of Comte [7, Proposition 7] for MCAR processes to MCARMA processes.

**Theorem 3.2.** *Let  $Y$  be a solution of the differential equation*

$$P(D)Y(t) = Q(D)DL(t), \quad t \geq 0,$$

with  $Y(0) = 0$  and polynomials  $P, Q$  given as in (1.3) and (1.4), respectively. Let the following assumptions hold:

- (a) If  $\det(P(z)) = 0$  then either  $\Re(z) < 0$  or  $z = 0$ .
- (b)  $\text{rank}(P_p) = \text{rank}(P(0)) = r$ ,  $0 < r < d$ , and  $P_p = \alpha\beta^\top$ , where the adjustment matrix  $\alpha \in M_{d,r}(\mathbb{R})$  and the cointegration matrix  $\beta \in M_{d,r}(\mathbb{R})$  have full rank  $r$ .
- (c)  $\alpha^\perp P_{p-1} \beta^\perp$  is of dimension  $(d-r) \times (d-r)$  with full rank  $(d-r)$ .

Then  $DY$  and  $\beta^\top Y$  are stationary processes, and thus, the MCARMA process  $Y$  is cointegrated.

*Proof.* By multiplying (3.1) with  $\alpha$  and  $\alpha^\perp$  we obtain with  $P_p = \alpha\beta^\top$  and  $\alpha^\perp \alpha = 0_{(d-r) \times r}$  the following equations

$$\begin{aligned} \alpha^\top Q(D)DL(t) &= \alpha^\top \alpha \beta^\top Y(t) + \alpha^\top P^*(D)DY(t), \\ \alpha^\perp Q(D)DL(t) &= \alpha^\perp P^*(D)DY(t). \end{aligned} \quad (3.2)$$

Since the system (3.2) is not invertible in  $Y$  and  $DY$  we define new processes

$$Z(t) := (\beta^\top \beta)^{-1} \beta^\top Y(t) \quad \text{and} \quad V(t) := (\beta^\perp{}^\top \beta^\perp)^{-1} \beta^\perp{}^\top DY(t) \quad \text{for } t \geq 0$$

and obtain thereby invertibility. The matrix  $R := (\beta, \beta^\perp) \in M_{d,d}(\mathbb{R})$  of rank  $d$  satisfies

$$R(R^\top R)^{-1} R^\top = \beta(\beta^\top \beta)^{-1} \beta^\top + \beta^\perp(\beta^\perp{}^\top \beta^\perp)^{-1} \beta^\perp{}^\top = I_d \quad (3.3)$$

since it is the sum of the projection matrices on the range and the projection matrix on the null space of  $\beta$ . Moreover, for  $\bar{\beta} := \beta(\beta^\top \beta)^{-1} \in M_{d,r}(\mathbb{R})$  and  $\bar{\beta}^\perp := \beta^\perp(\beta^\perp{}^\top \beta^\perp)^{-1} \in M_{d,d-r}(\mathbb{R})$  we have due to (3.3) that  $\beta \bar{\beta}^\top + \beta^\perp \bar{\beta}^\perp{}^\top = I_d$  holds. Furthermore, we have

$$DY(t) = (\beta \bar{\beta}^\top + \beta^\perp \bar{\beta}^\perp{}^\top) DY(t) = \beta DZ(t) + \beta^\perp V(t).$$

Rewriting system (3.2) with the newly defined variables yields

$$\begin{aligned} \alpha^\top Q(D) DL(t) &= \alpha^\top \alpha (\beta^\top \beta)^{-1} \beta^\top Z(t) + \alpha^\top P^*(D) \beta DZ(t) + \alpha^\top P^*(D) \beta^\perp V(t), \\ \alpha^\perp{}^\top Q(D) DL(t) &= \alpha^\perp{}^\top P^*(D) \beta DZ(t) + \alpha^\perp{}^\top P^*(D) \beta^\perp V(t). \end{aligned}$$

Rearranging the last expressions leads to

$$\tilde{P}(D)(Z(t)^\top, V(t)^\top)^\top = (\alpha, \alpha^\perp)^\top Q(D) DL(t),$$

where the matrix polynomial  $\tilde{P}$  is given by

$$\tilde{P}(z) := \begin{pmatrix} \alpha^\top \alpha (\beta^\top \beta)^{-1} + \alpha^\top P^*(z) \beta z & \alpha^\top P^*(z) \beta^\perp \\ \alpha^\perp{}^\top P^*(z) \beta z & \alpha^\perp{}^\top P^*(z) \beta^\perp \end{pmatrix}. \quad (3.4)$$

By assumption (b) and (c) we have

$$\det(\tilde{P}(0)) = \det \begin{pmatrix} \alpha^\top \alpha (\beta^\top \beta)^{-1} & \alpha^\top P^*(0) \beta^\perp \\ 0_{(d-r) \times r} & \alpha^\perp{}^\top P^*(0) \beta^\perp \end{pmatrix} = \det(\alpha^\top \alpha) \det(\beta^\top \beta) \det(\alpha^\perp{}^\top P^*(0) \beta^\perp) \neq 0,$$

where all matrices in the last line have full rank and consequently a non-zero determinant. Then we can see due to (3.1) and (3.4) that  $\tilde{P}(z) = (\alpha, \alpha^\perp)^\top P(z) (\beta, \beta^\perp / z)$  for  $z \neq 0$  and thus,

$$\det(\tilde{P}(z)) = \frac{1}{z^{d-r}} \det(\alpha, \alpha^\perp)^\top \det(P(z)) \det(\beta, \beta^\perp) \neq 0.$$

Thus,  $\tilde{P}(z)$  has the same roots as  $P(z)$ , except the null ones and the non-zero roots are assumed to have negative real part due to (a). Hence, the process  $(Z, V)$  is asymptotically (exponentially) stable and has a stationary solution. The process  $DY(t) = \beta DZ(t) + \beta^\perp V(t)$  is also stationary as a linear combination of stationary processes. Besides  $\beta^\top Y(t) = (\beta^\top \beta) Z(t)$  holds and therefore stationarity of  $\beta^\top Y(t)$  follows. Thus, the MCARMA process  $Y$  is cointegrated and this completes the proof.  $\square$

Finally, we make some remarks on the last result and its implications on cointegration for MCARMA models.

**Remark 3.3.** *The assumption in Theorem 3.2 have the following relevance:*

- Assumption (a) guarantees that the process is non-stationary.
- Assumption (b) guarantees that there exist linear combinations which are stationary.

- Assumption (c) guarantees that the process is integrated of order one and not of higher order.

Theorem 3.2 complements Lemma 2.6 which says that (a) and (b) are sufficient criteria for cointegration under the assumption that the state-space model is minimal. Conversely, Theorem 3.2 does not assume the minimality but instead (c).

**Remark 3.4.** If the cointegration rank is zero, i.e.  $P_p = 0_{d \times d}$ , we have no cointegration vector and thus, the process is not cointegrated. However, the process is integrated. On the other hand, if the rank of  $P_p$  is equal to  $d$ , i.e.  $P_p$  is of full rank, the process is stationary. This means that all eigenvalues have strictly negative real part and (a) is satisfied. Additionally, (c) is automatically satisfied, whereas (b) is violated. Therefore cointegration arises when the rank of  $P_p$  satisfies  $0 < r < d$ . Hence, it depends on the matrix  $P_p$  if we have a stationary, integrated or even cointegrated MCARMA process.

**Remark 3.5.** There are two natural ways to define an integrated MCARMA process. Both ways have the property that the process remains in the class of MCARMA processes. This can be seen by taking a closer look to the defining differential equations.

- (i) The first method starts with a stationary  $d$ -dimensional MCARMA( $p, q$ ) process  $Y$ . The integrated process is defined by integration of  $Y$ , namely  $I(t) = \int_0^t Y(s) ds$ . Assume that the process  $Y$  satisfies  $P(D)Y(t) = Q(D)DL(t)$  and define  $P^*(z) := zP(z)$ . Then the equation for the integrated process is

$$P^*(D)I(t) = P(D)DI(t) = Q(D)DL(t).$$

The order of the polynomial  $P^*(z)$  is  $p^* := p + 1$ . Obviously  $(I(t))_{t \geq 0}$  is then an MCARMA process as well with parameters  $(p^*, q)$  and  $p^* > q$ .

- (ii) The second method uses a non-stationary  $d$ -dimensional MCARMA( $p, q$ ) process  $I := (I(t))_{t \geq 0}$  where  $DI$  is stationary. Assume, that the process  $I$  satisfies  $P(D)I(t) = Q(D)DL(t)$ ,  $t \geq 0$ , and define  $Q^*(z) := zQ(z)$ . Then we have

$$P(D)DI(t) = D[P(D)I(t)] = D[Q(D)DL(t)] = Q^*(D)DL(t).$$

Clearly,  $DI(t) =: Y(t)$  is an MCARMA( $p, q + 1$ ) process. Again, this implies that we need the assumption  $p > q + 1$ .

The different definitions of integrated processes are not equivalent. Both have in common that  $DI$  is stationary and it is an MCARMA process, whereas in the first definition there exist no  $\beta$  so that  $\beta^\top I$  is stationary. Due to the different definition of integration in (ii),  $P_p$  is not fixed to be zero, thus we allow the process to be cointegrated.

### 3.2 Properties of cointegrated state-space models

Due to the decoupled canonical form the covariance matrix of the cointegrated state-space model can also be decomposed. We assume for reasons of simplicity that  $t_0 = 0$  for the rest of the paper.

**Corollary 3.6.** Let the cointegrated linear state-space model be defined by  $(A_2, B_1, B_2, C_1, C_2, L)$ . Then  $\mathbb{E}[Y(t)] = C_1 \mathbb{E}[X_1(0)]$  for  $t \geq 0$ . Suppose  $X_1(0) = 0$ . Then for  $t, s \geq 0$  we have

$$\text{Cov}(Y(t), Y(t+s)) = C_2 \exp(A_2 s) \Gamma_0 C_2^\top + \int_0^t C_2 \exp(A_2 u) B_2 \Sigma_L (C_1 B_1)^\top du$$

$$+ \int_s^{t+s} C_1 B_1 \Sigma_L B_2^\top \exp(A_2^\top u) C_2^\top du + t \cdot C_1 B_1 \Sigma_L (C_1 B_1)^\top,$$

where

$$\Gamma_0 = \int_0^\infty \exp(A_2 u) B_2 \Sigma_L (\exp(A_2 u) B_2)^\top du.$$

*Proof.* We obtain for the expectation evidently

$$\mathbb{E}[Y(t)] = \mathbb{E} \left[ C_1 X_1(0) + C_1 B_1 L(t) + C_2 \int_{-\infty}^t \exp(A_2(t-u)) B_2 dL(u) \right] = C_1 \mathbb{E}[X_1(0)].$$

Setting  $X_1(0) = 0$  leads then to

$$\begin{aligned} \text{Cov}(Y(t), Y(t+s)) &= \mathbb{E} \left[ Y(t) Y(t+s)^\top \right] \\ &= \mathbb{E} \left[ C_1 B_1 L(t) (C_1 B_1 L(t+s))^\top + C_1 B_1 L(t) Y_2(t+s)^\top + Y_2(t) (C_1 B_1 L(t+s))^\top + Y_2(t) Y_2(t+s)^\top \right] \\ &= C_1 B_1 \mathbb{E} \left[ L(t) L(t+s)^\top \right] (C_1 B_1)^\top + C_1 B_1 \mathbb{E} \left[ \int_0^t 1 dL(u) \left( \int_{-\infty}^{t+s} \exp(A_2(t+s-u)) B_2 dL(u) \right)^\top \right] C_2^\top \\ &\quad + C_2 \mathbb{E} \left[ \int_{-\infty}^t \exp(A_2(t-u)) B_2 dL(u) \left( \int_0^{t+s} 1 dL(u) \right)^\top \right] (C_1 B_1)^\top + \mathbb{E}[Y_2(t) Y_2(t+s)^\top] \\ &= C_1 B_1 \mathbb{E} \left[ L(t) L(t+s)^\top \right] (C_1 B_1)^\top + C_1 B_1 \mathbb{E} \left[ \int_0^t 1 dL(v) \left( \int_0^{t+s} \exp(A_2(t+s-u)) B_2 dL(u) \right)^\top \right] C_2^\top \\ &\quad + \mathbb{E}[Y_2(t) Y_2(t+s)^\top] + C_2 \mathbb{E} \left[ \int_0^t \exp(A_2(t-u)) B_2 dL(u) \left( \int_0^{t+s} 1 dL(v) \right)^\top \right] (C_1 B_1)^\top, \end{aligned}$$

and finally, the result follows by calculating all the remaining expectations using Marquardt and Stelzer [22, Proposition 3.13].  $\square$

The time dependence of the covariance function is clearly visible in this representation and hence, this process is indeed non-stationary.

### 3.3 Cointegrated state-space models sampled at a discrete time-grid

In this subsection, we study the sampled version of the continuous time cointegrated state-space model and derive the same decoupling (cf. Theorem 2.4). Note that the (co-)integration property of the continuous-time model directly transfers to its sampled version by Comte [7, Proposition 3].

**Corollary 3.7.** *Let the cointegrated linear state-space model be defined by  $(A_2, B_1, B_2, C_1, C_2, L)$ . Then the sampled process  $Y^{(h)} := (Y(nh))_{n \in \mathbb{N}}$  has the state-space representation*

$$\begin{pmatrix} X_{n,1}^{(h)} \\ X_{n,2}^{(h)} \end{pmatrix} = \begin{pmatrix} X_{n-1,1}^{(h)} \\ e^{A_2 h} X_{n-1,2}^{(h)} \end{pmatrix} + \begin{pmatrix} R_{n,1}^{(h)} \\ R_{n,2}^{(h)} \end{pmatrix}$$

and the observation equation

$$Y_n^{(h)} = C_1 X_{n,1}^{(h)} + C_2 X_{n,2}^{(h)},$$

with noise term

$$R_n^{(h)} = \begin{pmatrix} R_{n,1}^{(h)} \\ R_{n,2}^{(h)} \end{pmatrix} = \begin{pmatrix} B_1 (L(nh) - L((n-1)h)) \\ \int_{(n-1)h}^{nh} e^{A_2(nh-u)} B_2 dL(u) \end{pmatrix}, \quad n \in \mathbb{N}. \quad (3.5)$$

The sequence  $(R_n^{(h)})_{n \in \mathbb{N}}$  is an i.i.d. sequence with mean zero and covariance matrix

$$\tilde{\Sigma}^{(h)} := \begin{pmatrix} \tilde{\Sigma}_{11}^{(h)} & \tilde{\Sigma}_{12}^{(h)} \\ \tilde{\Sigma}_{21}^{(h)} & \tilde{\Sigma}_{22}^{(h)} \end{pmatrix} := \mathbb{E}(R_n^{(h)} R_n^{(h)\top}) = \int_0^h \begin{pmatrix} B_1 \Sigma_L B_1^\top & e^{A_2 u} B_2 \Sigma_L B_1^\top \\ B_1 \Sigma_L B_2^\top e^{A_2^\top u} & e^{A_2 u} B_2 \Sigma_L B_2^\top e^{A_2^\top u} \end{pmatrix} du.$$

Furthermore, we have that  $R_n^{(h)}$  has finite  $r^h$ -moment for some  $r > 0$  if the driving Lévy process  $L$  has finite  $r^h$ -moment. This implies the existence of the  $r^h$ -moment of  $Y^{(h)}$ .

*Proof.* The state-space representation follows at once by setting  $t = nh$  in (2.1) and the same holds for the covariance matrix. The existence of the  $r^h$ -moment follows immediately by Marquardt and Stelzer [22, Proposition 3.30] and Schlemm and Stelzer [28, Lemma 3.15].  $\square$

The corollary shows that we have also a separation of the stationary part and the integrated part in the linear state-space model of the discrete-time model as in the continuous-time model. Moreover, we can clearly see the connection of the eigenvalue zero of the transition matrix  $A$  with unit roots in the discrete time case since  $e^A$  has eigenvalues equal to one if  $A$  has eigenvalues equal to zero.

Let us now consider the solution of the sampled process  $Y^{(h)}$  in more detail. We have to replace on the one hand, the Lévy process by a random walk and on the other hand, the stationary continuous-time state-space model with its discrete-time version.

**Lemma 3.8.** *The solution of the sampled process  $Y^{(h)}$  given in Corollary 3.7 is*

$$Y_n^{(h)} = C_1 X_1^{(h)}(0) + C_1 B_1 L(nh) + Y_{n,2}^{(h)}, \quad n \in \mathbb{N}, \quad (3.6)$$

where the stationary part  $Y_2^{(h)} := (Y_{n,2}^{(h)})_{n \in \mathbb{N}}$  is

$$Y_{n,2}^{(h)} = C_2 \int_{-\infty}^{nh} e^{A_2(nh-u)} B_2 dL(u), \quad n \in \mathbb{N}. \quad (3.7)$$

*Proof.* In the same manner this follows by inserting  $t = nh$  into (2.2) and (2.4).  $\square$

The first two moments of the sampled process are derived in the next lemma.

**Lemma 3.9.** *The expectation of the sampled cointegrated state-space model (3.6) is given by*

$$\mathbb{E}[Y_n^{(h)}] = C_1 \mathbb{E}[X_1^{(h)}(0)], \quad n \in \mathbb{N}.$$

Suppose  $X_1^{(h)}(0) = 0$ . Then for  $n \in \mathbb{N}$ ,  $s \in \mathbb{N}_0$ , we have

$$\text{Cov}(Y_n^{(h)}, Y_{n+s}^{(h)}) = C_2 \exp(A_2 s) \tilde{\Sigma}_{2,2}^{(h)} C_2^\top + \int_0^{nh} C_2 \exp(A_2 u) B_2 \Sigma_L (C_1 B_1)^\top du$$



$$+ \int_{nh}^{(n+s)h} C_1 B_1 \Sigma_L B_2^\top \exp(A_2^\top u) C_2^\top du + nh \cdot C_1 \widetilde{\Sigma}_{1,1}^{(h)} C_1^\top.$$

*Proof.* Setting  $t = nh$  in Corollary 3.6 proves the claim.  $\square$

For the last part of this paper we need the first difference of the sampled cointegrated process and the stationary part of this process.

**Lemma 3.10.** *Let  $(R_{n,1}^{(h)})_{n \in \mathbb{N}}$  be defined as in (3.5). The first difference of the sampled cointegrated process  $Y^{(h)}$  is then*

$$\begin{aligned} \Delta Y_n^{(h)} &= C_1 R_{n,1}^{(h)} + \Delta Y_{n,2}^{(h)} \\ &= C_1 R_{n,1}^{(h)} + C_2 \int_{(n-1)h}^{nh} e^{A_2(nh-u)} B_2 dL(u) + C_2 \left( e^{A_2 h} - I_{N-c} \right) \int_{-\infty}^{(n-1)h} e^{A_2((n-1)h-u)} B_2 dL(u). \end{aligned} \quad (3.8)$$

*Proof.* The first equality is obvious due to (3.5) and (3.6). To show the last representation we use (3.7) which gives

$$\begin{aligned} \Delta Y_n^{(h)} &= C_1 R_{n,1}^{(h)} + \Delta Y_{n,2}^{(h)} \\ &= C_1 R_{n,1}^{(h)} + C_2 \int_{-\infty}^{nh} e^{A_2(nh-u)} B_2 dL(u) - C_2 \int_{-\infty}^{(n-1)h} e^{A_2((n-1)h-u)} B_2 dL(u) \\ &= C_1 R_{n,1}^{(h)} + C_2 \int_{(n-1)h}^{nh} e^{A_2(nh-u)} B_2 dL(u) + C_2 \left( e^{A_2 h} - I_{N-c} \right) \int_{-\infty}^{(n-1)h} e^{A_2((n-1)h-u)} B_2 dL(u) \end{aligned}$$

and consequently the statement.  $\square$

Note that the first difference of the sampled cointegrated process  $\Delta Y^{(h)}$  is obviously stationary and the  $r^{th}$ -moment exists if the  $r^{th}$ -moment of the Lévy process exists since  $Y^{(h)}$  has then a finite  $r^{th}$ -moment due to Corollary 3.7. Moreover, we know that  $(C_1 R_{n,1}^{(h)})_{n \in \mathbb{N}}$  is obviously an i.i.d. sequence.

Lastly, we analyze the dependence structure of the stationary part of the (sampled) process. Hence, we give the definition of strongly mixing; a kind of asymptotic independence. For more details on mixing processes see for example Bradley [3] or Doukhan [8].

**Definition 3.11.** *A continuous-time stationary stochastic process  $X = (X_t)_{t \geq 0}$  is called strongly mixing if for any  $m \in \mathbb{N}$*

$$\alpha_l := \sup \left\{ |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{F}_0^m, B \in \mathcal{F}_{m+l}^\infty \right\} \xrightarrow{l \rightarrow \infty} 0,$$

where  $\mathcal{F}_0^m := \sigma(X_t : 0 \leq t \leq m)$  and  $\mathcal{F}_{m+l}^\infty := \sigma(X_t : t \geq m+l)$ .

With regard to already known results, we see quite easily that the stationary part of the sampled process satisfies the following mixing property, and due to the decoupling we know then the dependence structure of each summand.

**Lemma 3.12.** *The stationary part  $Y_2$  of the state-space model  $Y$  defined by  $(A_2, B_1, B_2, C_1, C_2, L)$  is exponentially strongly mixing and the same holds true for the stationary part  $Y_2^{(h)}$  of the sampled*

process  $Y^{(h)}$ . There exists a constant  $\delta > 0$  such that the mixing coefficient  $\alpha_{Y_2^{(h)}}$  of the sampled process satisfies

$$\sum_{l=0}^{\infty} \left[ \alpha_{Y_2^{(h)}}(l) \right]^{\frac{\delta}{2+\delta}} < \infty.$$

The same holds true for  $C_1^\perp Y = C_1^\perp Y_2$  and  $C_1^\perp Y^{(h)} = C_1^\perp Y_2^{(h)}$ .

*Proof.* Due to Marquardt and Stelzer [22, Proposition 3.34], the assertion for the stationary process holds. This property transfers to the sampled process right away and we also have the condition on the mixing coefficients. The last claim follows directly by Bradley [3, Remark 1.8 b)].  $\square$

Furthermore, note that  $\Delta Y_2^{(h)}$  is also strongly mixing since it is the difference and consequently, a measurable function of the finite past values of a strongly mixing process (cf. [3, Remark 1.8 b)]).

## 4 Error correction form of a cointegrated state-space model sampled at a discrete time-grid

In this section we assume that we have a cointegrated state-space model  $(A_2, B_1, B_2, C_1, C_2, L)$ . In order to estimate the model parameters of the state-space model  $Y$  from observations in discrete time, the sampled version of the cointegrated state-space model is of special interest; more details can be found in Fasen and Scholz [12]. The representation of the discrete-time observations given in Corollary 3.7 has its limits since the noise  $R^{(h)}$  is not observable. For this reason we derive an alternative representation with the help of the linear innovations.

**Definition 4.1.** The linear innovations  $\varepsilon^{(h)} = (\varepsilon_n^{(h)})_{n \in \mathbb{N}}$  of  $Y^{(h)}$  are defined by

$$\varepsilon_n^{(h)} = Y_n^{(h)} - P_{n-1} Y_n^{(h)},$$

where  $P_n$  is the orthogonal projection onto  $\overline{\text{span}}\{Y_i^{(h)} : -\infty < i \leq n\}$  and the closure is taken in the Hilbert space of square-integrable random variables with inner product  $(X, Y) \mapsto \mathbb{E}\langle X, Y \rangle$ .

We want to obtain an error correction form of the sampled process. However, the linear state-space model is not in innovation form and thus, we apply a linear filter to this model. We use the Kalman filter for this purpose. The applicability of the Kalman filter for the sampled cointegrated state-space model given in Corollary 3.7 can be easily checked by adapting the results in Chui and Chen [6, Chapter 6] to the case of unit roots. The complete calculations can be found in Scholz [29, Section 4.6]. We only have to assume that the Lévy process has a positive definite covariance matrix, is independent of  $X_1(0)$  and that the matrix  $C$  has full rank  $d$  which is automatically satisfied by our cointegrated model. This implies  $\tilde{C}\tilde{\Sigma}^{(h)}C^\top$  has full rank and thus, the initialization of the Kalman filter with  $\Omega_0^{(h)} = \tilde{C}\tilde{\Sigma}^{(h)}C^\top$  guarantees the applicability of the Kalman filter.

Let us first sum up the important results concerning the Kalman filter in the next proposition.

**Proposition 4.2.** The discrete-time algebraic Riccati equation

$$\Omega^{(h)} := e^{Ah}\Omega^{(h)}e^{A^\top h} - e^{Ah}\Omega^{(h)}C^\top(C\Omega^{(h)}C^\top)^{-1}C\Omega^{(h)}e^{A^\top h} + \tilde{\Sigma}^{(h)} \in M_N(\mathbb{R})$$

has a positive definite solution  $\Omega^{(h)}$  and the steady state Kalman gain matrix  $K^{(h)}$  is given by

$$K^{(h)} := e^{Ah} \Omega^{(h)} C^T (C \Omega^{(h)} C^T)^{-1} \in M_{N,d}(\mathbb{R}).$$

The linear innovations  $\varepsilon^{(h)}$  of  $Y^{(h)}$  are the unique stationary solution of the linear state-space model

$$\begin{aligned} \widehat{X}_n^{(h)} &= (e^{Ah} - K^{(h)}C) \widehat{X}_{n-1}^{(h)} + K^{(h)} Y_{n-1}^{(h)}, \\ \varepsilon_n^{(h)} &= Y_n^{(h)} - C \widehat{X}_n^{(h)}, \quad n \in \mathbb{N}. \end{aligned} \quad (4.1)$$

Then, we have a moving average representation for the linear innovations given by

$$\begin{aligned} \varepsilon_n^{(h)} &= \left( I_d - C [I_N - (e^{Ah} - K^{(h)}C)B]^{-1} K^{(h)} B \right) Y_n^{(h)} \\ &= Y_n^{(h)} - C \sum_{i=1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)} Y_{n-i}^{(h)}, \end{aligned}$$

where  $B$  denotes the backshift operator defined as  $BY_n^{(h)} = Y_{n-1}^{(h)}$ . The covariance matrix of the innovations is given by

$$V^{(h)} = \mathbb{E} \left[ \varepsilon_n^{(h)} \varepsilon_n^{(h)T} \right] = C \Omega^{(h)} C^T \in \mathbb{S}_d^{++}.$$

Thus, the innovations form of  $Y^{(h)}$  is

$$\begin{aligned} \widehat{X}_n^{(h)} &= e^{Ah} \widehat{X}_{n-1}^{(h)} + K^{(h)} \varepsilon_{n-1}^{(h)}, \\ Y_n^{(h)} &= C \widehat{X}_n^{(h)} + \varepsilon_n^{(h)}, \quad n \in \mathbb{N}. \end{aligned} \quad (4.2)$$

The matrix exponential  $e^{Ah} = \exp(\text{diag}(0_{c \times c}, A_2 h)) = \text{diag}(I_c, \exp(A_2 h))$  where  $|\lambda_{\max}(\exp(A_2))| < 1$  by Theorem 2.4. Hence, (4.2) is indeed a cointegrated linear state-space model in discrete time.

Define now the rational matrix-valued transfer function

$$k(z) := I_d - C [I_N - (e^{Ah} - K^{(h)}C)z]^{-1} K^{(h)} z = I_d - C \sum_{i=1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)} z^i \quad \text{for } z \in \mathbb{C}. \quad (4.3)$$

Obviously, we have  $k(0) = I_d$  and

$$k(1) = I_d - C [I_N - (e^{Ah} - K^{(h)}C)]^{-1} K^{(h)} \in M_d(\mathbb{R}). \quad (4.4)$$

Due to the results known about the Kalman filter we know that  $|\lambda_{\max}(e^{Ah} - K^{(h)}C)| < 1$  (cf. for the cointegrated setting [29, Lemma 4.6.7] and the result in the stationary case can be found in Chui and Chen, [6, Lemma 6.8]), hence  $I_N - (e^{Ah} - K^{(h)}C)$  is invertible and  $k(1)$  is well defined. We want to show that  $k(1)$  contains the information of cointegration, in particular that  $k(1)$  has rank  $r = d - c$  and that the cointegration space can be determined with the knowledge of  $k(1)$ . In order to analyze the matrix  $k(1)$  we need the next assumption guaranteeing that the filtered system (4.2) is still minimal.

**Assumption B.** The linear model (4.2) is controllable, i.e.  $\mathcal{C}_{e^{Ah}K^{(h)}}$  has rank  $N$ .

Moreover, the linear model (4.2) is also observable due to the fact that the discrete-time and the continuous-time observability matrix coincide. Finally, Assumption B and the observability of the

model implies minimality (cf. Hannan and Deistler [15, Theorem 2.3.3]). The filtered minimal linear state-space model (4.2) is obviously given in decoupled form

$$\begin{pmatrix} \widehat{X}_{n+1,1}^{(h)} \\ \widehat{X}_{n+1,2}^{(h)} \end{pmatrix} = \begin{pmatrix} I_c & 0_{c \times (N-c)} \\ 0_{(N-c) \times c} & e^{A_2 h} \end{pmatrix} \begin{pmatrix} \widehat{X}_{n,1}^{(h)} \\ \widehat{X}_{n,2}^{(h)} \end{pmatrix} + \begin{pmatrix} K_1^{(h)} \\ K_2^{(h)} \end{pmatrix} \varepsilon_n^{(h)}, \quad (4.5)$$

$$Y_n^{(h)} = (C_1 \quad C_2) \begin{pmatrix} \widehat{X}_{n,1}^{(h)} \\ \widehat{X}_{n,2}^{(h)} \end{pmatrix} + \varepsilon_n^{(h)}.$$

We adapt a result from Ribarits and Hanzon, [24, Lemma 3.1] in order to characterize the rank of  $k(1)$ .

**Lemma 4.3.** *Suppose that the linear state-space model (4.2) satisfies Assumption B and has the representation (4.5). Further, let  $r$  denote the number of linearly independent cointegration relations of the corresponding process  $Y^{(h)}$ . Then we have*

- (i)  $r = d - c$  and  $C_1^\perp$  spans the cointegration space,
- (ii)  $\text{rank } k(1) = \text{rank} \left( I_d - C[I_N - (e^{A_1 h} - K^{(h)}C)]^{-1} K^{(h)} \right) = r$ .

*Proof.* (i) Follows directly from Lemma 2.6 and the minimality assumption.

- (ii) We obtain for  $k(1)$  the following representation by applying the decoupling into subsystems to (4.4)

$$k(1) = I_d - (C_1 \quad C_2) \begin{pmatrix} K_1^{(h)} C_1 & K_1^{(h)} C_2 \\ K_2^{(h)} C_1 & K_2^{(h)} C_2 + I_{N-c} - e^{A_2 h} \end{pmatrix}^{-1} \begin{pmatrix} K_1^{(h)} \\ K_2^{(h)} \end{pmatrix}.$$

Since  $K_1^{(h)}$  and  $C_1$  have full rank  $c$ , the  $c \times c$  matrix  $K_1^{(h)} C_1$  is regular and has also rank  $c$ . We set  $\tilde{N} := N - c$ . Furthermore, as before we know that the Kalman filter implies  $|\lambda_{\max}(e^{A_2 h} - K_2^{(h)} C_2)| < 1$  and consequently the matrix  $(K_2^{(h)} C_2 + I_{\tilde{N}} - e^{A_2 h})$  is nonsingular. Thus, we can apply the Matrix Inversion Lemma (see e.g. Bernstein [2], Proposition 2.8.7) and obtain

$$k(1) = I_d - (C_1 \quad C_2) \cdot M \cdot \begin{pmatrix} K_1^{(h)} \\ K_2^{(h)} \end{pmatrix}, \quad (4.6)$$

where the matrix  $M$  is defined by

$$M := \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad (4.7)$$

and

$$\begin{aligned} M_{11} &:= (K_1^{(h)} C_1)^{-1} + (K_1^{(h)} C_1)^{-1} K_1^{(h)} C_2 Q^{-1} K_2^{(h)} C_1 (K_1^{(h)} C_1)^{-1}, \\ M_{12} &:= -(K_1^{(h)} C_1)^{-1} K_1^{(h)} C_2 Q^{-1}, \\ M_{21} &:= -Q^{-1} K_2^{(h)} C_1 (K_1^{(h)} C_1)^{-1}, \\ M_{22} &:= Q^{-1}, \end{aligned}$$

with  $Q := I_{\tilde{N}} - e^{A_2 h} + K_2^{(h)} C_2 - K_2^{(h)} C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} C_2$ . Define  $P := I_d - C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} \in M_d(\mathbb{R})$ , which is obviously idempotent since  $P^2 = P$  holds. Note that the matrix product  $K_1^{(h)} C_1$  is a nonsingular  $c \times c$  matrix and consequently  $C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)}$  has rank  $c$ . Then we can rewrite the matrix  $k(1)$  once more using the representation (4.7) and obtain

$$\begin{aligned}
k(1) &= I_d - (C_1 \ C_2) \cdot M \cdot \begin{pmatrix} K_1^{(h)} \\ K_2^{(h)} \end{pmatrix} \\
&= I_d - C_1 M_{11} K_1^{(h)} - C_2 M_{21} K_1^{(h)} - C_1 M_{12} K_2^{(h)} - C_2 M_{22} K_2^{(h)} \\
&= (I_d - C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)}) - C_2 Q^{-1} K_2^{(h)} + C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} C_2 Q^{-1} K_2^{(h)} \\
&\quad + C_2 Q^{-1} K_2^{(h)} C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} - C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} C_2 Q^{-1} K_2^{(h)} C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)} \\
&= P - (I_d - C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)}) C_2 Q^{-1} K_2^{(h)} (I_d - C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)}) \\
&= P - P C_2 Q^{-1} K_2^{(h)} P.
\end{aligned}$$

Since the matrix  $P$  is idempotent and  $I - P = C_1 (K_1^{(h)} C_1)^{-1} K_1^{(h)}$  has obviously rank  $c$ , we have due to the rank equation for an idempotent matrix (i.e.  $\text{rank } P + \text{rank}(I - P) = d$ , see e.g. Bernstein [2, Fact 3.12.9]), that  $\text{rank } P = d - c$ . As above, by the Matrix Inversion Lemma we can rewrite the matrix  $Q^{-1}$  as

$$\begin{aligned}
Q^{-1} &= [I_{\tilde{N}} - e^{A_2 h} + K_2^{(h)} P^2 C_2]^{-1} \\
&= (I_{\tilde{N}} - e^{A_2 h})^{-1} - (I_{\tilde{N}} - e^{A_2 h})^{-1} K_2^{(h)} P [I_d + P C_2 (I_{\tilde{N}} - e^{A_2 h})^{-1} K_2^{(h)} P]^{-1} P C_2 (I_{\tilde{N}} - e^{A_2 h})^{-1}.
\end{aligned}$$

For the sake of brevity, we write  $R := C_2 (I_{\tilde{N}} - e^{A_2 h})^{-1} K_2^{(h)}$ . Substituting the previous result into the formula for  $k(1)$  leads to

$$\begin{aligned}
k(1) &= P - PRP + PRP(I_d + PRP)^{-1} PRP = P - PRP + (PRP)^2 (I_d + PRP)^{-1} \\
&= [(P - PRP)(P + PRP) + (PRP)^2] (I_d + PRP)^{-1} = P(I_d + PRP)^{-1},
\end{aligned}$$

where we used the fact that  $(I_d + AB)^{-1} A = A(I_d + BA)^{-1}$  for matrices  $A$  ( $:= PRP$ ) and  $B$  ( $:= I_d$ ) such that  $I_d + AB$  is nonsingular (see e.g. [2, Fact 2.16.16]) for the second equality. We can conclude

$$\text{rank } k(1) = \text{rank } P(I_d + PRP)^{-1} = \text{rank } P = d - c = r.$$

Thus, we have completed the proof.  $\square$

Since  $\text{rank } k(1) = r = d - c$  there exists  $\alpha, \beta \in M_{d,r}(\mathbb{R})$  with full row rank such that  $k(1) = -\alpha\beta^\top$ . Define

$$\tilde{k}(z) := I_d - \frac{k(z) - k(1)z}{1 - z} \quad (4.8)$$

which can be represented as an infinite order linear filter due to  $k(z) - k(1)z = 0$  for  $z = 1$  and  $k(z) - k(1)z = I_d$  for  $z = 0$ . Hence, we can rewrite  $k(z)$  as

$$k(z) = k(1)z + [k(z) - k(1)z] = k(1)z + (1 - z)[I_d - \tilde{k}(z)]$$

$$= k(1)z + I_d(1-z) - \tilde{k}(z)(1-z) = -\alpha\beta^\top z + I_d(1-z) - \tilde{k}(z)(1-z).$$

This leads to

$$\begin{aligned}\varepsilon_n^{(h)} &= k(B)Y_n^{(h)} = \left[ -\alpha\beta^\top B + I_d(1-B) - \tilde{k}(B)(1-B) \right] Y_n^{(h)}, \\ &= -\alpha\beta^\top Y_{n-1}^{(h)} + \Delta Y_n^{(h)} - \tilde{k}(B)\Delta Y_n^{(h)}, \quad n \in \mathbb{N}.\end{aligned}$$

We can now state an error correction form where we consider a linear state-space model instead of a VAR process as in the classical error correction form. The so-called transfer function error correction form for discrete-time state-space models was presented by Ribarits and Hanzon [24]. However, we have a continuous-time state-space model observed at discrete time points. Therefore, our error correction form has a similar form as the one of Ribarits and Hanzon but we have different matrices.

**Definition 4.4.** *The error correction form is given by*

$$\Delta Y_n^{(h)} = \alpha\beta^\top Y_{n-1}^{(h)} + \tilde{k}(B)\Delta Y_n^{(h)} + \varepsilon_n^{(h)}, \quad n \in \mathbb{N}. \quad (4.9)$$

For comparison see the classical error correction model for a cointegrated VARMA process for example in Lütkepohl [21, Section 14.2]. Having this error correction form enables us to analyze the short-run and long-run behavior of the process separately; we directly see the cointegration space in the representation.

**Lemma 4.5.** *The process  $\beta^\top Y^{(h)}$  is stationary and the rows of  $\beta$  span the cointegration space.*

*Proof.* By the Matrix Inversion Lemma (see e.g. [2, Corollary 2.8.8]) and (4.6) we obtain

$$\begin{aligned}k(1) &= I_d - (C_1 \ C_2) \cdot \begin{pmatrix} K_1^{(h)}C_1 & K_1^{(h)}C_2 \\ K_2^{(h)}C_1 & K_2^{(h)}C_2 + I_{N-c} - e^{A_2h} \end{pmatrix}^{-1} \cdot \begin{pmatrix} K_1^{(h)} \\ K_2^{(h)} \end{pmatrix} \\ &= [I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}] - [I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}]C_1 \\ &\quad \cdot (K_1^{(h)}[I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}]C_1)^{-1}K_1^{(h)}[I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}].\end{aligned}$$

We receive

$$k(1)C_1 = [I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}]C_1 - [I_d - C_2(I_{N-c} - e^{A_2h} + K_2^{(h)}C_2)^{-1}K_2^{(h)}]C_1 = 0_{d \times c}.$$

This means  $-\alpha\beta^\top C_1 = k(1)C_1 = 0_{d \times c}$ . Since  $\alpha$  and  $\beta$  have full rank  $r = d - c$  and  $\text{rank } C_1^\perp = d - c$ ,  $C_1^\perp$  and  $\beta$  span the same space. Due to Lemma 2.6 we can conclude the statement.  $\square$

In particular  $\varepsilon_n^{(h)}$  is then as a sum of stationary processes stationary itself. In the following we present some alternative representations for the innovation sequence and present some probabilistic properties of them.

**Lemma 4.6.** *Write  $k(z) = \sum_{j=1}^{\infty} L_j z^j$  and  $\tilde{k}(z) = \sum_{j=1}^{\infty} \tilde{K}_j z^j$ , and define  $\bar{k}(z) := I_d - \tilde{k}(z)$ . Then the following alternative representations for the innovation sequence hold:*

$$\begin{aligned}(a) \quad \varepsilon_n^{(h)} &= k(B)Y_{n,2}^{(h)} + \bar{k}(B)C_1 R_{n,1}^{(h)}, \quad n \in \mathbb{N}. \\ (b) \quad \varepsilon_n^{(h)} &= \sum_{j=0}^{\infty} \left( \tilde{K}_j C_1 B^j - \sum_{k=0}^j L_k C_2 e^{A_2h(j-k)} B^j \right) R_n^{(h)}, \quad n \in \mathbb{N}, \text{ where} \\ &\quad \tilde{K}_0 = 0 \text{ and } \tilde{K}_j = C \sum_{i=j+1}^{\infty} (e^{A_2h} - K^{(h)}C)^{i-1} K^{(h)}, \quad \text{for } j \geq 1.\end{aligned}$$



From (a) we see that the innovations consist of two summands containing on the one hand, the filtered version of the stationary part and on the other hand, the filtered version of the increments of the driving Lévy process.

*Proof.*

(a) Using (4.8) we obtain

$$k(z) - k(1)z = (1 - z)[I_d - \tilde{k}(z)] = (1 - z)\bar{k}(z)$$

so that (4.9) and (3.8) result in

$$\begin{aligned}\varepsilon_n^{(h)} &= \Delta Y_n^{(h)} + k(1)BY_n^{(h)} - \tilde{k}(B)\Delta Y_n^{(h)} = k(1)BY_{n,2}^{(h)} + \bar{k}(B)\Delta Y_n^{(h)} \\ &= k(1)BY_{n,2}^{(h)} + \bar{k}(B)\left(C_1R_{n,1}^{(h)} + (1 - B)Y_{n,2}^{(h)}\right) \\ &= k(B)Y_{n,2}^{(h)} + \bar{k}(B)C_1R_{n,1}^{(h)}, \quad n \in \mathbb{N}.\end{aligned}$$

(b) With the moving average representation of  $Y_2^{(h)}$  and the Cauchy product we receive

$$\begin{aligned}\varepsilon_n^{(h)} &= \bar{k}(B)C_1R_{n,1}^{(h)} + k(B)Y_{n,2}^{(h)} \\ &= \sum_{j=0}^{\infty} \tilde{K}_j B^j C_1 R_{n-j,1}^{(h)} + \sum_{j=0}^{\infty} L_j B^j C_2 \sum_{i=0}^{\infty} e^{A_2 h i} B^i R_{n,2}^{(h)} \\ &= \sum_{j=0}^{\infty} \tilde{K}_j B^j C_1 R_{n,1}^{(h)} + \sum_{i=0}^{\infty} \sum_{j=0}^k L_j B^j C_2 e^{A_2 h(i-j)} B^{i-j} R_{n,2}^{(h)} \\ &= \sum_{j=0}^{\infty} \tilde{K}_j B^j C_1 R_{n,1}^{(h)} + \sum_{j=0}^{\infty} \sum_{k=0}^j L_k C_2 e^{A_2 h(j-k)} B^j R_{n,2}^{(h)} \\ &= \sum_{j=0}^{\infty} \left( \tilde{K}_j C_1 B^j \quad \sum_{k=0}^j L_k C_2 e^{A_2 h(j-k)} B^j \right) \begin{pmatrix} R_{n,1}^{(h)} \\ R_{n,2}^{(h)} \end{pmatrix} \\ &= \sum_{j=0}^{\infty} \left( \tilde{K}_j C_1 B^j \quad \sum_{k=0}^j L_k C_2 e^{A_2 h(j-k)} B^j \right) R_n^{(h)}, \quad n \in \mathbb{N}.\end{aligned}$$

We determine the matrix coefficients  $\tilde{K}_i$  in the subsequent considerations. It can easily be seen that  $\tilde{k}(0) = \tilde{K}_0 = 0$ . By rearranging (4.8) we obtain  $k(z) - k(1)z = (1 - z)[I_d - \tilde{k}(z)]$  which is equivalent to

$$(1 - z) \left[ I_d - \sum_{i=1}^{\infty} \tilde{K}_i z^i \right] = \left( I_d - C \sum_{i=1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)} z^i \right) - \left( I_d - C \sum_{i=1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)} \right) z$$

and thus, by comparison of the coefficients we obtain

$$\begin{aligned}z^1 : \quad & -CK^{(h)} - I_d + C \sum_{i=1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)} = -I_d - \tilde{K}_1 \Rightarrow \tilde{K}_1 = C \sum_{i=2}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)}, \\ z^2 : \quad & -C(e^{Ah} - K^{(h)}C)K^{(h)} = \tilde{K}_2 - \tilde{K}_1 \Rightarrow \tilde{K}_2 = C \sum_{i=3}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1} K^{(h)}, \\ & \vdots\end{aligned}$$

$$z^j : -C(e^{Ah} - K^{(h)}C)^{j-1}K^{(h)} = \tilde{K}_j - \tilde{K}_{j-1} \Rightarrow \tilde{K}_j = C \sum_{i=j+1}^{\infty} (e^{Ah} - K^{(h)}C)^{i-1}K^{(h)}, \quad \text{for } j \geq 1.$$

This concludes the proof.  $\square$

In the end, we investigate useful properties of the linear innovation sequence in the next proposition.

**Proposition 4.7.** *The linear innovations  $\varepsilon^{(h)}$  given in (4.1) are a stationary, ergodic and uncorrelated sequence. Furthermore, the linear innovations  $\varepsilon^{(h)}$  have finite second moments.*

*Proof.* As already mentioned  $\varepsilon^{(h)}$  is stationary because all other terms appearing in (4.9) are stationary. For the ergodic property we first note  $R_1^{(h)}$  is an i.i.d. sequence and  $Y_2^{(h)}$  is ergodic, which was already shown in Schlemm and Stelzer [28]. Then, the vector process

$$(Z_n^{(h)})_{n \in \mathbb{N}} = \begin{pmatrix} C_1 R_{n,1}^{(h)} \\ Y_{n,2}^{(h)} \end{pmatrix}_{n \in \mathbb{N}}$$

is obviously ergodic and stationary. Since by Lemma 4.6 (a), we can define a measurable function  $f$  such that  $\varepsilon_n^{(h)} = f(Z_n^{(h)})$ ,  $n \in \mathbb{N}$ , and obtain that  $\varepsilon^{(h)}$  is ergodic with Bradley [3], Proposition 2.10 (ii).

Due to the already used result  $|\lambda_{\max}(e^{Ah} - K^{(h)}C)| < 1$  (cf. [29, Lemma 4.6.7]) and the form of the transfer function (4.3), the transfer function  $k(z)$  has exponentially decaying coefficients. Thus, the existence of the second moment of  $\varepsilon_n^{(h)} = k(B)Y_n^{(h)}$  follows directly from the finite second moment of the driving Lévy process.  $\square$

Note that if the Lévy process is a Brownian motion then  $Y^{(h)}$  and hence,  $\varepsilon^{(h)}$  is Gaussian. This even implies that the linear innovations are a sequence of i.i.d. random variables.

## References

- [1] Bauer, D. and Wagner, M. (2012). A state space canonical form for unit root processes. *Econometric Theory* **28**, 1313–1349.
- [2] Bernstein, D. S. (2009). *Matrix mathematics: theory, facts, and formulas* 2nd ed. ed. Princeton Univ. Press, Princeton.
- [3] Bradley, R. C. (2007). *Introduction to strong mixing conditions* vol. 1. Kendrick, Heber City.
- [4] Brenner, R. J. and Kroner, K. F. (1995). Arbitrage, cointegration, and testing the unbiasedness hypothesis in financial markets. *J. Financ. Quant. Anal.* **10**, 23–42.
- [5] Chambers, M. J. (1999). Discrete time representation of stationary and non-stationary continuous time systems. *J. Econ. Dyn. Control* **23**, 619–639.
- [6] Chui, C. K. and Chen, G. (2009). *Kalman filtering: with real-time applications* 4th ed. Springer, Berlin.
- [7] Comte, F. (1999). Discrete and continuous time cointegration. *J. Econometrics* **88**, 207–226.
- [8] Doukhan, P. (1994). *Mixing: properties and examples*. Lecture Notes in Statist. 85. Springer, New York.
- [9] Engle, R. F. and Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica* **55**, 251–276.
- [10] Fasen, V. (2013). Time series regression on integrated continuous-time processes with heavy and light tails. *Econometric Theory* **29**, 28–67.

- [11] Fasen, V. (2014). Limit theory for high frequency sampled MCARMA models. *Adv. Appl. Probab.* **46**, 846–877.
- [12] Fasen-Hartmann, V. and Scholz, M. (2016). Estimation of cointegrated linear state-space models in continuous time. In preparation.
- [13] Gantmacher, F. R. (1986). *Matrizentheorie*. Springer, Berlin , Heidelberg [u.a.].
- [14] Granger, C. (1981). Some properties of time series data and their use in econometric model specification. *J. Econometrics* **16**, 121–130.
- [15] Hannan, E. and Deistler, M. (2012). *The Statistical Theory of Linear Systems*. Society for Industrial and Applied Mathematics, Philadelphia.
- [16] Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* **59**, 1551–1580.
- [17] Johansen, S. (2009). *Likelihood-based inference in cointegrated vector autoregressive models*. Advanced texts in econometrics. Oxford Univ. Press, Oxford.
- [18] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.* 35–45.
- [19] Kessler, M. and Rahbek, A. (2001). Asymptotic likelihood based inference for co-integrated homogenous Gaussian diffusions. *Scand. J. Statist.* **28**, 455–470.
- [20] Kessler, M. and Rahbek, A. (2004). Identification and inference for multivariate cointegrated and ergodic Gaussian diffusion. *Stat. Inference. Stoch. Process.* **7**, 137–151.
- [21] Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer, Berlin.
- [22] Marquardt, T. and Stelzer, R. (2007). Multivariate CARMA processes. *Stochastic Process. Appl.* **117**, 96–120.
- [23] Phillips, P. C. B. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica* **59**, 967–980.
- [24] Ribarits, T. and Hanzon, B. (2011). The state-space error correction model: Definition, estimation and model selection. *Unpublished manuscript received from the authors upon request*.
- [25] Sargan, J. D. (1964). Wages and prices in the united kingdom: a study in econometric methodology. *Econometric analysis for national economic planning* **16**, 25–54.
- [26] Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge Stud. Adv. Math. 68. Cambridge Univ. Press, Cambridge.
- [27] Schlemm, E. and Stelzer, R. (2012). Multivariate CARMA processes, continuous-time state space models and complete regularity of the innovations of the sampled processes. *Bernoulli* **18**, 46–63.
- [28] SCHLEMM, E. and Stelzer, R. (2012). Quasi maximum likelihood estimation for strongly mixing state space models and multivariate Lévy-driven CARMA processes. *Electron. J. Stat.* **6**, 2185–2234.
- [29] Scholz, M. (2016). Estimation of Cointegrated Multivariate Continuous-time Autoregressive Moving Average Processes. *PhD thesis*. Karlsruher Institute of Technology, Karlsruhe.
- [30] Shieh, L. S., Tsay, Y. T. and Yates, R. (1983). Block transformations of a class of multivariable control systems to four basic block companion forms. *Comput. Math. Appl.* **9**, 703 – 714.
- [31] Stockmarr, A. and Jacobsen, M. (1994). Gaussian diffusions and autoregressive processes: weak convergence and statistical inference. *Scand. J. Statist.* **21**, 403–429.